

## PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2002-297588

(43)Date of publication of application : 11.10.2002

(51)Int.Cl.

G06F 17/28

(21)Application number : 2001-102266

(71)Applicant : FUJITSU LTD

(22)Date of filing : 30.03.2001

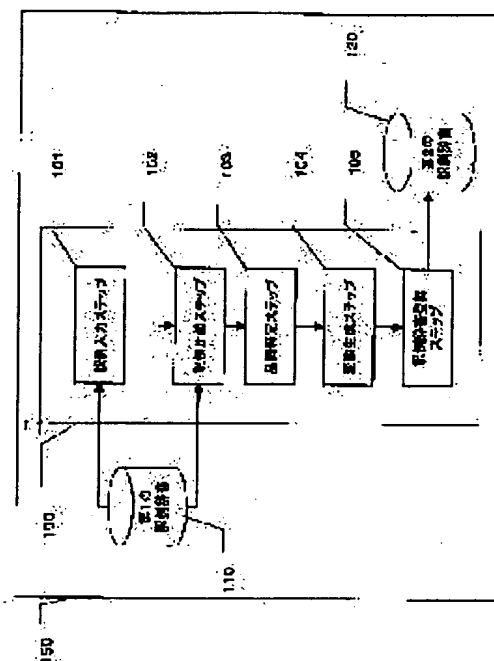
(72)Inventor : OZAWA HITOSHI  
KAWAZU TAKAKO

## (54) METHOD, PROGRAM AND APPARATUS FOR AUTOMATIC CREATION FOR DICTIONARY OF TRANSLATION EXAMPLE

## (57)Abstract:

**PROBLEM TO BE SOLVED:** To automatically create a dictionary of translation example where a part of translations is parameterized.

**SOLUTION:** A creation method for dictionary of translation example to use translation of translation example comprises, a translation example comparison step 102 for comparing a first translation example with other first translation example, a part of speech specifying step 103 for specifying a part of speech of a site of difference detected by the step 102, a parameter creation step 104 for creating a second translation example parameterized with liking the site detected by the step 102 with the part of speech specified by the step 103 and a registration step 105 for dictionary of translation example to register the second example into the dictionary.



## LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

Copyright (C); 1998,2003 Japan Patent Office

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号  
特開2002-297588  
(P2002-297588A)

(43) 公開日 平成14年10月11日 (2002. 10. 11)

(51) Int.Cl.<sup>7</sup>  
G 0 6 F 17/28

識別記号

F I  
G 0 6 F 17/28

テーマコード\* (参考)  
C 5 B 0 9 1

審査請求 未請求 請求項の数 5 O L (全 9 頁)

(21) 出願番号 特願2001-102266 (P2001-102266)

(22) 出願日 平成13年3月30日 (2001. 3. 30)

(71) 出願人 000005223  
富士通株式会社  
神奈川県川崎市中原区上小田中4丁目1番  
1号  
(72) 発明者 小澤 仁  
東京都大田区新蒲田一丁目17番25号 株式  
会社富士通ラーニングメディア内  
(72) 発明者 河津 香子  
東京都大田区新蒲田一丁目17番25号 株式  
会社富士通ラーニングメディア内  
(74) 代理人 100108187  
弁理士 横山 淳一

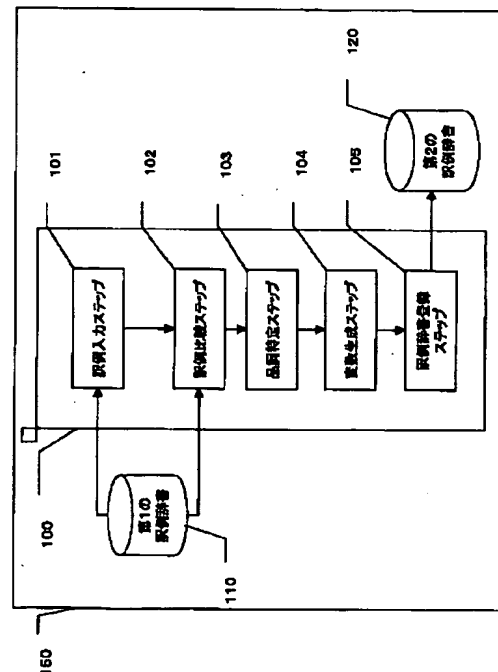
最終頁に続く

(54) 【発明の名称】 訳例辞書の自動生成方法およびプログラムおよび装置

(57) 【要約】 (修正有)

【課題】 訳文の一部が変数化された訳例辞書を、自動的に生成することを目的とする。

【解決手段】 訳例翻訳に用いる訳例辞書の作成方法において、第1の訳例と他の第1の訳例とを比較する訳例比較ステップ102と、訳例比較ステップで検出した差異の部位の品詞を特定する品詞特定ステップ103と、訳例比較ステップで検出した差異の部位と品詞特定ステップが特定した品詞とを関連づけて変数化した第2の訳例を作成する変数生成ステップ104と、第2の訳例を訳例辞書に登録する訳例辞書登録ステップ105とを含むことを特徴とする、訳例辞書を自動生成する方法によって解決される。



## 【特許請求の範囲】

【請求項 1】 訳例翻訳に用いる訳例辞書の作成方法において、

第 1 の訳例と他の第 1 の訳例とを比較する訳例比較ステップと、

前記訳例比較ステップで検出した差異の部位の品詞を特定する品詞特定ステップと、

前記訳例比較ステップで検出した差異の部位と前記品詞特定ステップが特定した品詞とを関連づけて変数化した第 2 の訳例を作成する変数生成ステップと、

前記第 2 の訳例を訳例辞書に登録する訳例辞書登録ステップとを含むことを特徴とする訳例辞書を自動生成する方法。

【請求項 2】 前記訳例辞書登録ステップは、前記変数生成ステップで変数化した部位がいくつあったかを記録することを特徴とする請求項 1 の訳例辞書を自動生成する方法。

【請求項 3】 訳例翻訳に用いる辞書の作成方法において、第 1 の訳例辞書から、訳例の一部を変数化した第 2 の訳例辞書を自動生成するプログラムであって、

第 1 の訳例と他の第 1 の訳例とを比較する訳例比較ステップと、

前記訳例比較ステップで検出した差異の部位の品詞を特定する品詞特定ステップと、

前記訳例比較ステップで検出した差異の部位と前記品詞特定ステップが特定した品詞とを関連づけて変数化した第 2 の訳例を作成する変数生成ステップと、

前記第 2 の訳例を訳例辞書に登録する訳例辞書登録ステップとを含むことを特徴とする訳例辞書を自動生成するプログラム。

【請求項 4】 訳例翻訳に用いる訳例辞書の作成装置において、

第 1 の訳例と他の第 1 の訳例とを比較する訳例比較手段と、

前記訳例比較ステップで検出した差異の部位の品詞を特定する品詞特定手段と、

前記訳例比較ステップで検出した差異の部位と前記品詞特定ステップが特定した品詞とを関連づけて変数化した第 2 の訳例を作成する変数生成手段と、前記第 2 の訳例を訳例辞書に登録する訳例辞書登録手段とを含むことを特徴とする訳例辞書を自動生成する装置。

【請求項 5】 訳例翻訳に用いる訳例辞書の作成するためのプログラムを格納した記録媒体であって、

第 1 の訳例と他の第 1 の訳例とを比較する訳例比較ステップと、

前記訳例比較ステップで検出した差異の部位の品詞を特定する品詞特定ステップと、

前記訳例比較ステップで検出した差異の部位と前記品詞特定ステップが特定した品詞とを関連づけて変数化した第 2 の訳例を作成する変数生成ステップと、

前記第 2 の訳例を訳例辞書に登録する訳例辞書登録ステップとを実行させるためのプログラムを記録したコンピュータ読み取り可能な記録媒体。

## 【発明の詳細な説明】

## 【0001】

【発明が属する技術分野】 本発明は、訳例の一部が変数化された訳例辞書の自動生成方法に関する。

## 【0002】

【従来の技術】 機械翻訳システムで、ある言語で書かれた原文を翻訳しようとしたとき、コンピュータがもつ基本辞書あるいは専門用語辞書による翻訳だけでは原文の意味に即した翻訳を行うことは難しい。そこで、翻訳担当者があらかじめ頻繁に出現すると思われる原文と訳文からなる訳例を機械翻訳システムの訳例辞書に登録することにより該当する原文を翻訳処理する実行する訳例翻訳という方法がある。

【0003】 訳例翻訳において使用する訳例辞書の様式として、訳例の一部、例えば名詞などを変数化したものがある。翻訳対象の原文が、訳例辞書の原文の変数を除いた部分と一致した場合に、原文と訳文の変数部分を、コンピュータがもつ基本辞書や専門用語辞書から補い翻訳を行う方法に用いる訳例辞書である。

【0004】 この訳例の一部を変数化して持つ訳例辞書を作成するにあたっては、翻訳者が原文と訳文のどの部位を変数化するかを判断し、手で登録を行う必要があるため多くの工数を必要とし、また作業により変数化する部位の基準が異なることがあった。

【0005】 また、変数化した部位が多い訳例辞書から翻訳された訳文は、変数化された部位が少ない訳例辞書から翻訳された訳文と比べると信頼性が低いものになるが、その判断を行う材料を提供できなかった。

## 【0006】

【発明が解決しようとする課題】 以上の問題に鑑み本発明は、訳例の一部が変数化された訳例辞書を、自動的に生成すること、訳例の中で変数化された部位の数を記録すること、を目的とするものである。

## 【0007】

【課題を解決するための手段】 第 1 の訳例と他の第 1 の訳例とを比較する訳例比較ステップと、前記訳例比較ステップで検出した第 1 の訳例の差異の部位を変数化した、第 2 の訳例を作成する変数生成ステップと、前記第 2 の訳例を訳例辞書に登録する訳例辞書登録ステップとを含むことを特徴とする、訳例辞書を自動生成する方法。

【0008】 を含む訳例辞書の自動生成方法によって解決される。

## 【0009】

【発明の実施の形態】 本発明の実施の形態を 図 1 と図 2 と図 3 に示す。

【0010】 訳例辞書の自動生成プログラム 100 は、

第1の訳例辞書110と、訳例入力ステップ101と、訳例比較ステップ102と、品詞特定ステップ103と、変数生成ステップ104と、訳例辞書登録ステップ105と、第2の訳例辞書120とを有するように構成する。

【0011】訳例辞書の自動生成装置150は、第1の訳例辞書110と、訳例入力ステップ101と、訳例比較ステップ102と、品詞特定ステップ103と、変数生成ステップ104と、訳例辞書登録ステップ105と、第2の訳例辞書120とが稼動するように構成する。

【0012】第1の訳例辞書110は、図2で示すように変数化された部位のない、原文201と訳文202からなる第1の訳例200を持つように構成する。

【0013】第2の訳例辞書120は、図3で示すように訳例の一部を変数化した、原文301と訳文302からなる第2の訳例300を持つように構成する。なお、原文301と訳文302以外に変数化した部位の数303を持ってよい。

【0014】訳例入力ステップ101は、第1の訳例辞書110から図2で示す原文201と訳文202からなる第1の訳例200を読み込むように構成する。

【0015】訳例比較ステップ102は、原文入力ステップ101が読み込んだ第1の訳例200の原文201と、第1の訳例辞書にある他の訳例の原文との差異部位を比較するように構成する。

【0016】品詞特定ステップ103は、訳例比較ステップ102において検出された差異とされる部位の品詞を特定するように構成する。

【0017】変数生成ステップ104は、訳例比較ステップ102において検出された差異とされる部位と品詞特定ステップ103において特定された品詞の種類とを関連付けて変数化した図3で示す原文301と訳文302を持つ第2の訳例を作成するように構成する。なお、変数生成ステップ104は、第2の訳例として原文と訳文以外にも変数の数303を作成してもよい。

【0018】訳例辞書登録ステップ105は、変数生成ステップ104が作成した図3で示す第2の訳例300を、第2の訳例辞書120に登録するように構成する。

【0019】本発明の一実施の形態を、図4から図10を用いて示す。

【0020】ここでは、元の訳例辞書401から比較元となる訳例610と、比較する対象となる訳例620とを読み込み、両者が近似の場合は差異の部位を変数化した第2の訳例630を新規の訳例辞書403に登録する場合を例に説明を行う。

【0021】比較元となる訳例と、比較対象となる訳例とを第1の訳例辞書から以下に示すステップS0からステップS4において読み込む。

【0022】初期値として、比較する両者の訳例が近似

か否かを判断する際のしきい値となる値 $x$ をユーザから受け付けるとともに、第1の訳例辞書から読み込む位置を示す変数 $m$ と $n$ と、差異の部位の数を判定する $y$ と、差異の部位を判定する処理に用いる変数 $L$ に初期値として0をセットする。(ステップS0)

元の訳例辞書401から読み込む、比較元となる訳例の位置 $m$ と、比較対象となる訳例の位置 $n$ を決める。(ステップS1)

元の訳例辞書401から $m$ 番目にある比較元となる訳例610を読み込む。(ステップS2)

比較元となる訳例610と比較する対象となる訳例の読み込み位置を $n+1$ 番目にする。(ステップS3)

元の訳例辞書401から $n$ 番目にある比較する対象となる訳例620を読み込む。(ステップS4)

読み込んだ両者の訳例の原文を品詞単位に複数の部位に分割し、比較することにより両者の訳例が近似であるか否かの判定を以下に示すステップS5からステップS10において行う。

【0023】ステップS2で読み込んだ訳例610の原文611を形態素解析し、品詞単位に複数の部位に分割する。(ステップS5)

S4で読み込んだ訳例620の原文621を形態素解析し、品詞単位に複数の部位に分割する。(ステップS6)

ステップS5とステップS6で分割した部位単位の比較位置を $L+1$ 番目にする。(ステップS7)

ステップS5とステップS6で分割した各々の $L$ 番目の部位の文字列が一致するかを比較する。なお、一致の判定としては意味の一致を用いてもよい。(ステップS8)

ステップS8で比較した結果が一致しない場合、差異の部位の数 $y$ を $+1$ する。(ステップS9)

差異の部位の数 $y$ を判定する。もし差異の部位 $y$ が $x$ 以上あれば比較する対象となる訳例620は比較元の訳例610とは近似ではないと判断し、差異の部位の数 $y$ と部位単位の比較位置 $L$ を0に設定した後、新たに比較する対象となる訳例620を読み込み比較対象訳例読込処理(S4)に遷移する。(ステップS10)

原文の差異の部位を変数化する処理を以下に示すステップS31からステップS32において行う。

【0024】差異の部位 $y$ が $x$ 未満なら、差異の部位の品詞の特定を行う。(ステップS31)

比較元となる訳例610の原文611と比較する対象となる訳例620の原文621の差異の部位と、ステップS31で特定した品詞とを関連付けて変数にする。(ステップS32)

訳文の差異の部位を変数化する処理を以下に示すステップS33からステップS38において行う。

【0025】比較する対象となる訳例620の訳文622を形態素解析し、品詞単位に複数の部位に分割する。

(ステップS33)

ステップS31で特定した品詞とステップS33で分割したJ番目の部位の品詞の種類とを比較する。(ステップS34)

もし両者の品詞の種類が一致しなければ、 $J = J + 1$ を行い次の部位の比較を行う。(ステップS35)

もし両者の品詞の種類が一致すれば、ステップS33で分割した部位のJ番目の部位を基本辞書または専門用語辞書などの一般の辞書402を用いて原文の言語に翻訳を行う。(ステップS36)

もし原文の差異の部位とステップS36の訳で一致する訳がなければ、 $J = J + 1$ を行い次の部位の比較を行う。(ステップS37)

もしステップS36で一致する訳があれば、比較する対象となる訳例620の訳文621のJ番目の部位と、J番目の部位の品詞とを関連付けて変数化する。(ステップS38)

訳例の一部を変数化した第2の訳例630を第2の訳例辞書に登録する処理を以下に示すステップS50において行う。

【0026】ステップS32とステップS38で作成した訳例630を、変数化した部位の数631と併せて、新規の訳例辞書403に登録する。なお、差異の部位の数が0個の場合は、新規の訳例辞書403への登録は行わないようにしてもよい。差異の部位を判定する処理に用いる変数Lの値を0に初期化する。(ステップS50)

なお、ここでは元の訳例辞書401と新規の訳例辞書402はおのおの1個の例で説明を行ったが、複数の元の訳例辞書401から複数の新規の訳例辞書402を作成する、あるいは複数の元の訳例辞書401から1個の新規の訳例辞書402を作成する、または1個の元の訳例辞書401から複数の新規の訳例辞書402を作成するようにしてもよい。

【0027】また、新規の訳例辞書402は新規ではなく、既存の訳例辞書に追加される形でよい。

【0028】変数化された部位の数631を用いる一実施の形態を、図11を用いて示す。変数化された部位の数を持つ訳例辞書710を用いて原文群720を翻訳し、訳文群730を作成するものとする。

【0029】原文(その1)721から、訳例辞書で変数化された部位が1箇所しかなかった訳例(その1)711を用いて作成された訳文(その1)731を緑色で表示する。

【0030】原文(その2)722から変数化された部位が2箇所あった訳例712(その2)を用いて作成された訳文(その2)732を黄色で表示する。

【0031】原文(その3)723から、変数化された部位が3箇所もあった訳例(その3)713を用いて作成された訳文(その3)733は赤色を表示する。

(付記1) 訳例翻訳に用いる訳例辞書の作成方法において、第1の訳例と他の第1の訳例とを比較する訳例比較ステップと、前記訳例比較ステップで検出した差異の部位の品詞を特定する品詞特定ステップと、前記訳例比較ステップで検出した差異の部位と前記品詞特定ステップが特定した品詞とを関連づけて変数化した第2の訳例を作成する変数生成ステップと、前記第2の訳例を訳例辞書に登録する訳例辞書登録ステップとを含むことを特徴とする訳例辞書を自動生成する方法。

10 (付記2) 前記訳例辞書登録ステップは、前記変数生成ステップで変数化した部位がいくつあったかを記録することを特徴とする付記1の訳例辞書を自動生成する方法。

(付記3) 訳例翻訳に用いる辞書の作成方法において、第1の訳例辞書から、訳例の一部を変数化した第2の訳例辞書を自動生成するプログラムであって、第1の訳例と他の第1の訳例とを比較する訳例比較ステップと、前記訳例比較ステップで検出した差異の部位の品詞を特定する品詞特定ステップと、前記訳例比較ステップで検出した差異の部位と前記品詞特定ステップが特定した品詞とを関連づけて変数化した第2の訳例を作成する変数生成ステップと、前記第2の訳例を訳例辞書に登録する訳例辞書登録ステップとを含むことを特徴とする訳例辞書を自動生成するプログラム。

(付記4) 訳例翻訳に用いる訳例辞書の作成装置において、第1の訳例と他の第1の訳例とを比較する訳例比較手段と、前記訳例比較ステップで検出した差異の部位の品詞を特定する品詞特定手段と、前記訳例比較ステップで検出した差異の部位と前記品詞特定ステップが特定した品詞とを関連づけて変数化した第2の訳例を作成する変数生成手段と、前記第2の訳例を訳例辞書に登録する訳例辞書登録手段とを含むことを特徴とする訳例辞書を自動生成する装置。

(付記5) 訳例翻訳に用いる訳例辞書の作成するためのプログラムを格納した記録媒体であって、第1の訳例と他の第1の訳例とを比較する訳例比較ステップと、前記訳例比較ステップで検出した差異の部位の品詞を特定する品詞特定ステップと、前記訳例比較ステップで検出した差異の部位と前記品詞特定ステップが特定した品詞とを関連づけて変数化した第2の訳例を作成する変数生成ステップと、前記第2の訳例を訳例辞書に登録する訳例辞書登録ステップとを実行させるためのプログラムを記録したコンピュータ読み取り可能な記録媒体。

【0032】

【発明の効果】本発明により、訳例の一部が変数化された訳例辞書の生成を自動的に行うとりわけ利用価値の高いものである。たとえば、既存の訳例辞書から、多くの工数を必要とせずに、また作業者により変数化する部位の基準が異なることなく訳例の一部が変数化された訳例辞書を自動的に生成することができる。

【0033】また、変数化された部位の少ない訳例から翻訳された比較的信頼性の高い訳文と、変数化された部位の多い訳例から翻訳された比較的信頼性の低い訳文とを区別して表示をする際の判断などに用いられるように、訳例の中で変数化した部位の数を記録することができる。

【図面の簡単な説明】

【図1】 原理を示す図

【図2】 第1の訳例辞書を示す図

【図3】 第2の訳例辞書を示す図

【図4】 処理フローを示す図（その1）

【図5】 処理フローを示す図（その2）

【図6】 処理フローを示す図（その3）

【図7】 処理フローを示す図（その4）

【図8】 処理フローを示す図（その5）

【図9】 変数生成を示す図

【図10】 第2の訳例の作成方法を示す図

【図11】 変数の数により翻訳結果を色分け表示することを示す図

【符号の説明】

100 訳例辞書の自動生成プログラム

101 訳例入力ステップ

102 訳例比較ステップ

103 品詞特定ステップ

10 104 変数生成ステップ

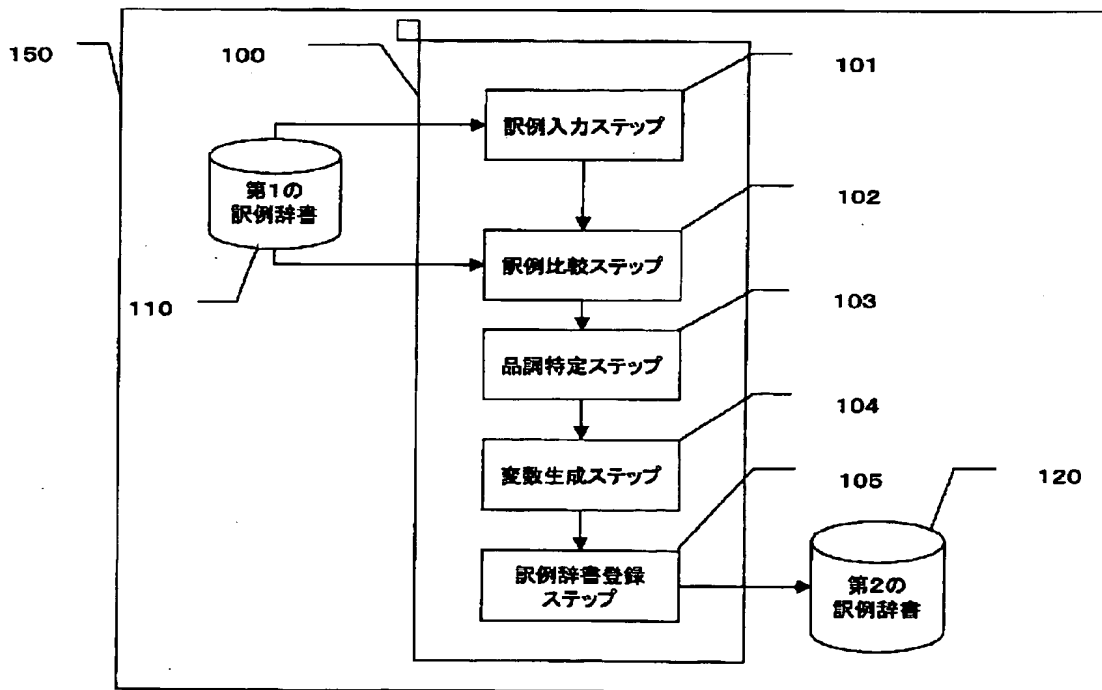
105 訳例辞書登録ステップ

110 第1の訳例辞書

120 第2の訳例辞書

150 訳例辞書の自動生成装置

【図1】



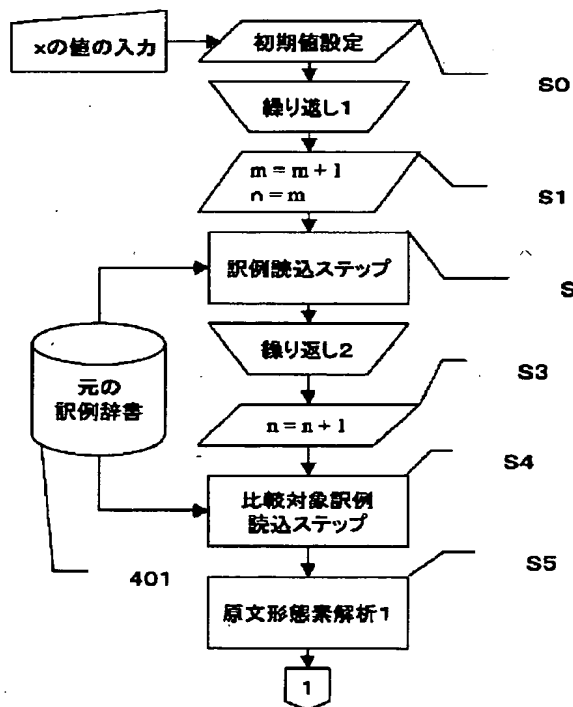
【図3】

301 原文	302 訳文	303 変数化した部位の数
<n1>は黒字。	They are black figures the <n1> half.	1
<n1>は<n2>だった。	<n1> were <n2>.	2
<n1>は<adv1>。	<n1> is <adv1>.	2
<n1>の<n2>は<n3>に転じた。	The <n2> of the <n1> were changed into the <n3>.	3

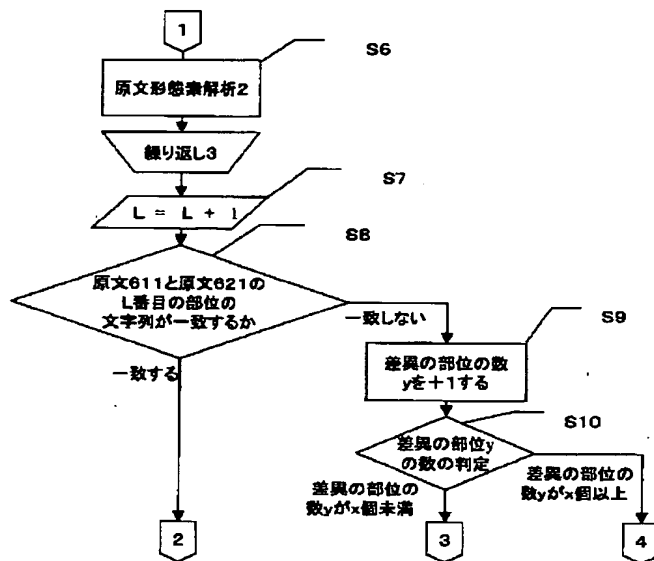
【図2】

201 原文	202 訳文
上期は黒字。	They are black figures for first half.
下期は黒字。	They are black figures the second half.
業績は黒字だった。	Achievements were black figures.
業績は赤字だった。	Achievements were red figures.
予算の達成は難しい。	Achievement of budget is difficult.
予算の達成はやさしい。	Achievement of budget is easy.
上期の業績は黒字に転じた。	The achievements of the first half were changed into black figures.
下期の業績は黒字に転じた。	The achievements of the second half were changed into black figures.

【図4】

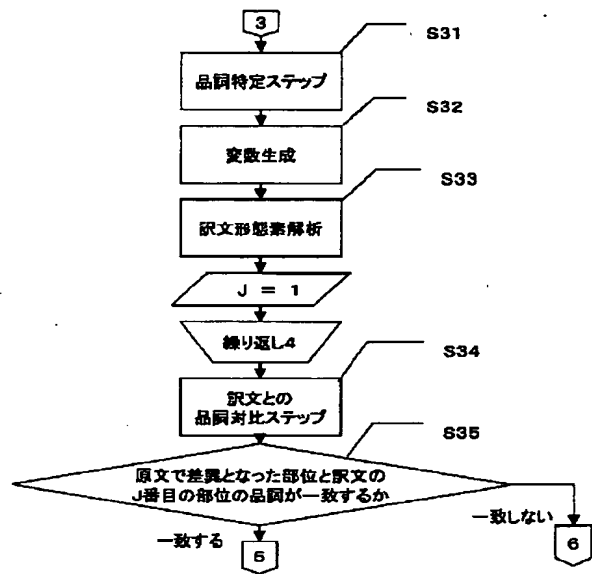


【図5】

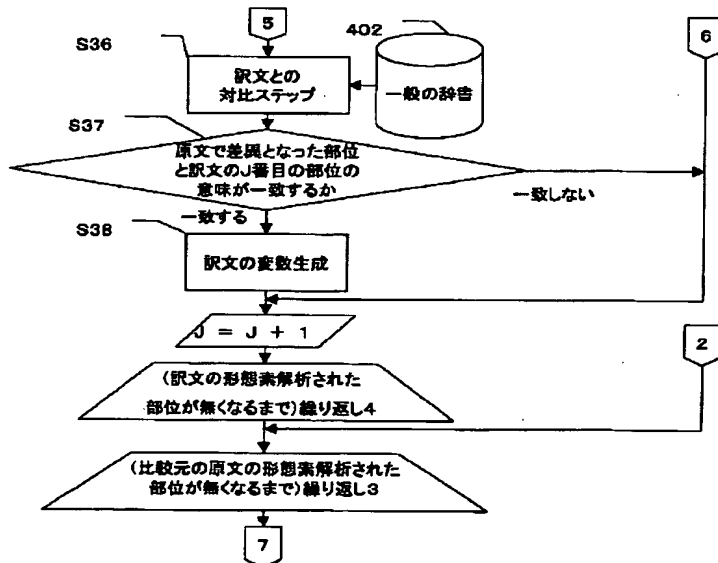




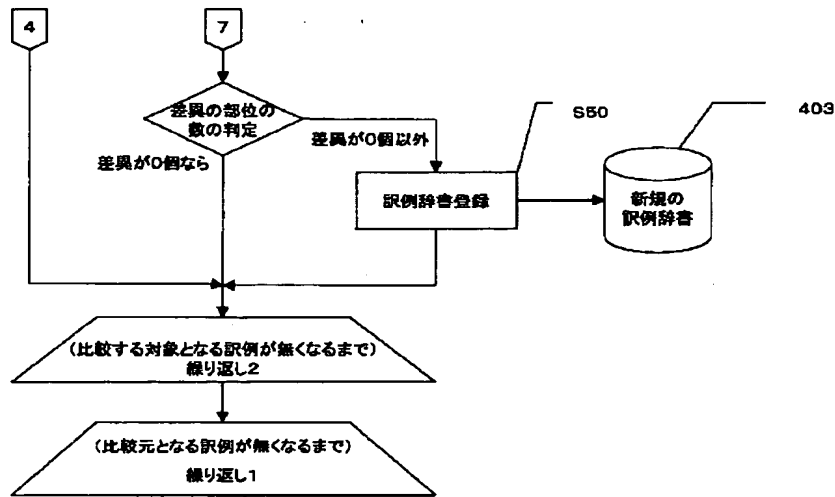
【図6】



【図7】



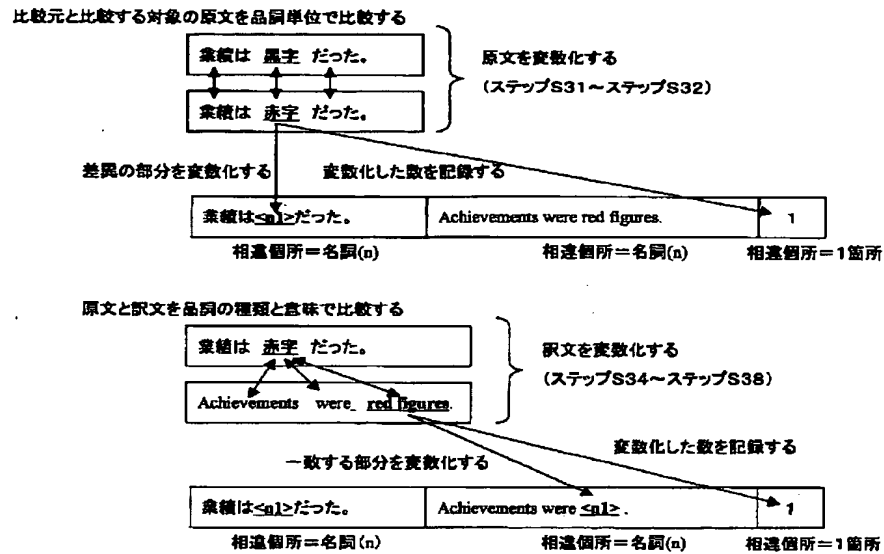
【図8】



【図9】

610	611 業績は黒字だった。	612 Achievements were black figures.
	621 業績は赤字だった。	622 Achievements were red figures.
620		
630	業績は<n1>だった。	Achievements were <n1>.
		631 1

【図10】



【図11】

原文	訳文	変数の数
<n1>です。	It is <n1>.	1
<n1>は<adv1>。	There are <adv1> <n1>.	2
<n1>の<n2>は<n3>だった。	The <n2> of the <n1> were <n3>.	3

原文
不況です。
売上は少ない。
下期の業績は赤字だった。
訳文
It is depression.
There are few sales.
The achievements of the second half were in the red.

フロントページの続き

Fターム(参考) 5B091 AA05 BA05 CC01 CC15 CC16